DOCUMENT RESUME

ED 291 886 CE 049 691

AUTHOR Whitney, Douglas R.; And Others
TITLE The Reliability and Validity of the GED Tests. GED Research Brief No. 6.
INSTITUTION American Council on Education, Washington, DC. GED Testing Service.
PUB DATE May 85
NOTE 5p.
PUB TYPE Information Analyses (070)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Adult Education; *High School Equivalency Programs; Item Analysis; Scores; Test Items; *Test Reliability; *Test Validity
IDENTIFIERS *General Educational Development Tests

ABSTRACT

This research brief summarizes the available reliability and validity data available in, but spread throughout, a number of General Educational Development (GED) Testing Service publications. A section on reliability discusses how to determine reliability of a test's scores and two ways of assessing the reliability of a test--internal consistency reliability and parallel forms reliability. The section on validity defines the term and considers some of the major elements of validity. These include content validity, concurrent validity, and predictive validity. A list of references is appended. (YLB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D. Whitney

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

AMERICAN COUNCIL ON EDUCATION ged℠ RESEARCH BRIEF

GED TESTING SERVICE OF THE AMERICAN COUNCIL ON EDUCATION • MAY 1985 • NO. 6

ED291886

The Reliability and Validity of the GED Tests

Douglas R. Whitney • Andrew G. Malizio • Wayne M. Patience

Much technical information has been published concerning the reliability and validity of the Tests of General Educational Development (GED Tests), but this information is spread throughout a number of GEDTS publications. To provide a convenient summary of this information for interested persons, we have briefly summarized the available reliability and validity data in this Research Brief. (Some information is published here for the first time.) For readers interested in a more extensive discussion, references to existing publications are given.

RELIABILITY

The reliability of a test's scores refers to their degree of *accuracy*. That is, to what extent would an examinee be expected to earn similar scores taking the same test again, taking different forms of the test and/or taking the test on different days? Reliability can be estimated either from a single test administration (internal consistency reliability) or from administering two forms of the test to a sample of examinees (parallel forms reliability). Reliability is also affected by the nature of the examinee population. For the GED Tests, the two populations of primary interest are the high school seniors in the norming studies and the actual GED examinees.

Internal Consistency Reliability. One way of assessing the reliability of a test is to compute a coefficient reflecting the internal consistency of the examinee's responses across the set of items in the test. The Kuder-Richardson formula 20 (KR20) coefficient is often used for this purpose. KR20 coefficients of .90 to .99 are usually obtained for tests that have a high degree of internal consistency—a desirable property in most educational tests. Because many different forms of each of the five tests in the GED battery are used, we have reported the ranges of KR20 coefficients that resulted from administering the tests in 1980 to samples of U.S. graduating high school seniors and to samples of GED examinees. Descriptions of the 1980 high school norming sample are found in Spille (1981a, 1981b) and the 1980 GED candidate sample is described by Malizio and Whitney (1981).

GED TEST	Range of KR20 coefficients (10 test forms) U.S. H.S. Seniors	Range of KR20 coefficients (9 test forms) U.S. GED Examinees
Writing Skills	.93 - .94	.88 - .94
Social Studies	.91 - .93	.86 - .94
Science	.90 - .93	.86 - .93
Reading Skills	.89 - .92	.85 - .93
Mathematics	.90 - .93	.81 - .92

Similar KR20 coefficients were obtained for seven recent test forms administered to samples of U.S. high school seniors in 1983 and 1985. These values would generally be considered acceptable evidence of a high degree of internal consistency (reliability), with the coefficients for seniors slightly higher than those for GED examinees. This result is chiefly due to the fact that the seniors' scores are somewhat more variable—have a larger standard deviation—than those for GED examinees. (See Malizio & Whitney, 1982, p. 4, Table 1.)

Another way to express the degree of internal consistency of a test is to report the "standard error of measurement" based on the internal consistency coefficients obtained. Based on the KR20 values above, the average standard errors of measurement (in GED standard score points) are reported below. These standard errors of measurement represent the range of score values within which examinees would typically be expected to score if they had taken the same form of the test, on the same day, under identical conditions (that is, within two or three standard score points).

GED TEST	Average Std. Error of Meas. (Standard Scores) (10 test forms) U.S. H.S. Seniors	Average Std. Error of Meas. (Standard Scores) (9 test forms) U.S. GED Examinees
Writing Skills	2.6	2.4
Social Studies	2.8	2.6
Science	2.9	2.7
Reading Skills	3.1	2.9
Mathematics	2.9	2.8

Parallel Forms Reliability. Another indication of the degree of reliability for tests is the correlation between the scores for examinees when two different forms of the test are taken

Douglas R. Whitney, Associate Director
Andrew G. Malizio, Senior Research Associate
Wayne M. Patience, Senior Research Associate
GED Testing Service
One Dupont Circle, NW
Washington, DC 20036

CE 049691

by the same persons. This coefficient is sometimes called a parallel forms or alternate forms reliability. During the 1980 norming study, high school seniors took two different forms of each of two GED Tests. These data permit us to calculate and report parallel form reliabilities for our high school sample. Because one would want to see a very close relationship between the scores a person would earn on different forms of the same test, coefficients between .80 and .95 are generally sought for educational tests. The parallel forms reliabilities based on the administration of the GED Tests to a national sample of high school seniors are reported below. The corresponding standard errors of measurement reflect the range of values within which examinees would be expected to score if they took different forms of the test on different days (that is, within three or four standard score points).

GED TEST	Range of Parallel forms reliabilities (10 test forms) U.S. High School Seniors	Average Std. Error of Measurement (Standard Scores)
Writing Skills	.84 - .89	3.6
Social Studies	.80 - .88	3.9
Science	.80 - .88	4.1
Reading Skills	.76 - .86	4.2
Mathematics	.82 - .86	3.9

These reliability values are slightly lower than the corresponding KR20 values for the high school examinees. This is because the parallel forms reliabilities are based on different forms of the tests administered on different days. These reliabilities are, however, still in the range generally considered to be satisfactory.

VALIDITY

The validity of a test refers to the *meaningfulness* of its scores. That is, to what degree are an examinee's scores interpretable as measures of the intended (or assumed) knowledge and skills? Because tests often have more than one purpose, and examinees have many reasons for taking the test, validity can be evaluated separately for each purpose and reason. Meaningful interpretation involves considering the tests' content, the correlations of the tests' scores with data from other sources (e.g., other tests, school grades), and the experiences of users of the tests' results (APA, 1974). Because there are many aspects of validity for the GED Tests, many pieces of information are relevant. Some of the major elements of validity are considered in the following sections.

Content Validity. The GED Tests are intended to measure the major and lasting outcomes of a high school program of study. Accordingly, the single most important aspect of validity for the GED Tests is their content validity. One way of assessing content validity for the GED Tests is to compare the tests to what is generally contained in the typical high school curriculum. As noted in an ETS report (1978, pp. 1-5), the content and test specifications for the current GED Tests were developed as a result of a 1974 meeting of subject matter specialists. The content of the current GED Tests is fully described in the publication *What do the GED Tests measure?* (Patience & Whitney, 1982, pp. 7-52). GEDTS is now updating the content for the tests to be introduced in 1988 based on a similar meeting held in February 1984.

There is no statistical method commonly employed to compute a coefficient of content validity; rather the establishment of this kind of validity is usually based on experts' subjective analysis of the test-curriculum similarities and differences. The GED Testing Service intends that the similarities be substantial, and accordingly involves high school and adult education curriculum specialists extensively in the development of the test specifications and the tests. As reported in a recent issue of the *GED Items* newsletter (GEDTS, 1984b), the items for a recent set of test forms were written by 45 experienced educators from around the country and reviewed by 19 other experienced educators. In addition to the careful and detailed descriptions of test content, the involvement of many additional educators in writing and reviewing the tests adds some assurance of content validity.

Ultimately, the degree of content validity inherent in the GED Tests must be judged by experts not directly involved in the development of the tests. Based on a study by Patience and Whitney (1983), such independent experts generally judge the current tests to reflect the outcomes of a high school program of study and to measure these outcomes fairly and adequately.

Concurrent Validity. Because the GED Tests are designed to determine whether or not adult examinees (non-high school graduates) have acquired a sufficient level of educational skills and knowledge to be granted a state or province high school equivalency credential, there should be a high degree of consistency between the GED score standards set by the states and provinces and those implicitly used by the high schools. To gather validity data on this issue, and in fact to ensure that such consistency does occur, the GED Testing Service regularly administers the GED Tests to national samples of high school graduating seniors. (See Spille, 1981a, 1981b.) These data permit the interpretation that, at the GED score levels required by most of the states, between 27% and 33% of the current graduating high school seniors would be unable to pass the GED Tests (GEDTS, 1984a, p. 63). Further information concerning the test results for seniors, all GED examinees, and those GED examinees who pass the tests appears in Malizio and Whitney (1983, p. 8, Table 15). These data confirm that the standards employed in the GED testing program are somewhat more stringent than those employed by the nation's high schools.

There are five tests in the GED battery—each designed to measure different content as indicated by the test titles. Because each of the tests requires that the examinee read and interpret written material, and because four of the tests involve the use of written passages followed by a series of questions, one should expect that the scores on the tests will be somewhat related. That is, examinees who earn high (or low) scores on one test will tend to earn high (or low) scores on the others. The degree to which this occurs is reflected in the correlations between scores on any pair of tests for examinees who have taken both. These correlations are reported below for 683 U.S. high school seniors and for approximately 11,000 GED examinees who completed all five tests as a part of the 1980 norming and GED examinee studies. (Correlations for seniors are above the diagonal; those for GED examinees are below the diagonal.) Although these correlations are substantial, each

is somewhat lower than the typical correlation between scores on different forms of the same tests (the parallel forms reliability). This suggests that the five tests in the GED battery measure strongly related, but somewhat distinct, skills and knowledge.

GED TEST	Writing Skills	Social Studies	Science	Reading Skills	Mathematics
	U.S. High School Seniors				
Writing Skills		80	76	76	69
Social Studies	74		82	79	71
Science	72	80		77	73
Reading Skills	74	79	75		69
Mathematics	63	65	66	60	
	U.S. GED Examinees				

As states have begun to require graduation tests for their high school seniors, another source of data relating to the standards employed for the GED Tests has become available. Sonnenblick (1980) reports evidence that resulted from the administration of New York's minimum competency graduation reading test to a small sample of GED examinees. In this study, 37% of those who passed the state's required test failed to pass the GED Tests; none of those who failed the state's test were able to pass the GED Tests. In a recent study in Florida, Stoker (1985) concluded that students who cannot pass the required Florida State Student Assessment Test, Part II are unlikely to succeed on the GED Tests.

Another aspect of concurrent validity for the GED Tests can be drawn from studies that report the correlations between scores on the GED Tests and those on other educational tests. Many examples are to be found in Spille's (1984) bibliography, but a few will be presented here. Sonnenblick (1980) reports that the correlation between average scores on the GED test battery and scores on New York's graduation reading test was .77. Correlations of the state test scores with the scores from individual GED Tests ranged from .60 to .67 for GED Tests 1-4 and .45 for the GED Mathematics test. (One would expect that GED Mathematics test scores would correlate less strongly with a reading test than would the scores on the other four tests in the GED battery.) Stoker (1985) reports that correlations between scores on Florida's SSAT-II (a criterion- referenced test used as Florida's high school graduation requirement) communications test and those on the GED Writing Skills and Reading Skills tests were .47 and .55 respectively. The correlation between scores on the Florida math test and on the GED Mathematics test was .58. For a sample of about 150 GED graduates who also took the ACT Assessment, the correlations between similarly-named tests in the two batteries ranged from .57 for the ACT and GED Social Studies tests to .71 for the ACT English and GED Writing Skills tests; ACT Composite scores correlated with GED average standard scores .74. (These correlations were between scores on tests taken up to two years apart.) Similarly, when the GED Tests are administered to high school graduating seniors, there is a positive correlation (.58) between GED average standard scores and student self-reported high school grade averages. Finally, a recent GEDTS study of the relationship between standard scores on the GED Writing Skills test and scores on holistically-scored essay tests found correlations ranging from .55 to .69. Generally, the correlations between scores on tests in the GED battery and those from similar tests in other batteries are substantial—although the range of correlations is wide and the magnitude of the correlations depends on the similarity of the tests, the time that elapses between test administrations, and differences in examinee motivation.

Predictive Validity. Another aspect of the GED Tests' validity is the degree to which the nation's employers and postsecondary institutions confer equivalent status to persons who have graduated from high school in the traditional way (and at the traditional age) with those who earned their equivalency credentials through the GED program. There is evidence from national surveys that businesses and higher education institutions accept the credentials based on successful performance on the GED Tests as equivalent to those earned in the more traditional fashion. For example, Spille (1982) reports that 95% of the nation's postsecondary education institutions consider applicants with high school equivalency credentials and those with traditional diplomas identically. In the business sector, Malizio and Whitney (1985) found that between 96% and 98% of the companies surveyed in a national study give GED graduates the same initial levels of employment, the same starting salary, and the same opportunities for advancement as are given to traditional high school graduates. A recent study of the acceptance of the GED credential in the hiring policies of Denver area employers indicated that 83% consider the GED credential and the high diploma to be equivalent for hiring purposes when all other factors are equal (Carson, 1983).

Yet another aspect of predictive validity is the degree to which the tests serve their intended purpose(s) for the examinees. The results of nationwide follow-up studies of GED graduates (Cervero & Peterson, 1982; Behal, 1983) suggest that many, though not all, of the examinees' expectations are met after passing the tests. Studies in Maryland (Reed, 1984) and New Jersey (Darkenwald & Valentine, 1984, 1985) have reached similar conclusions. For persons seeking greater employment and educational opportunities, passing the GED Tests clearly leads to such improvements for a majority of the examinees. GED graduates regularly report improvements in pay, acceptance in education and training programs, and other expected benefits.

Although the GED testing program is not limited to examinees who wish to attend postsecondary education institutions, the GED examinee population contains many persons interested in attending training or educational programs. It is estimated that 200,000 to 250,000 GED graduates annually enter some type of postsecondary education program. Once in these programs, the academic success of GED graduates (compared to that of traditional high school graduates) reflects the predictive validity of the GED Tests. A number of studies of the college success of GED graduates have been completed. (See, for example, Wolf, 1980; Wilson, *et al.*, 1981; Colert, 1983.) Although there are minor differences in the results of these studies, it can be generally concluded that GED graduates (as a group) succeed in college at about the same rate and to about the same degree as do traditional high school graduates.

There are other aspects of validity that are not addressed here, but the authors hope that these data and elaborations will be an adequate introductory summary of the available evidence. If additional or specific information is needed, please contact the authors.

REFERENCES

American Psychological Association, American Educational Research Association, & National Council on Measurement in Education. (1974). *Standards for educational and psychological tests*. Washington, DC: American Psychological Association.

Behal, E. L. (1983). *After the GED Tests: Postsecondary education enrollment patterns and perceptions of GED examinees*. (GEDTS Research Studies, No. 5). Washington, DC: American Council on Education.

Carson, B. W. (1983). *Acceptance of General Educational Development (GED) in hiring policies of Denver area employers*. (Unpublished doctoral dissertation, University of Northern Colorado.)

Cervero, R. M., & Peterson, N. L. (1982). *After the GED Tests: Employment and training activities of GED examinees*. (GEDTS Research Studies, No. 4). Washington, DC: American Council on Education.

Colert, S. (1983, December). *High school equivalency and high school diploma students at Brandon University: A comparison of academic success*. (GEDTS Research Brief No. 5). Washington, DC: American Council on Education.

Darkenwald, G. G., & Valentine, T. (1984). *Outcomes and impact of adult basic education*. (Research Monograph No. 6). New Brunswick, NJ: Center for Adult Development, Rutgers University.

Darkenwald, G. G., & Valentine, T. (1985, April). *The General Educational Development Tests: Benefits of graduation and a typology of graduates*. Paper presented at the meeting of the American Educational Research Association, Chicago.

Educational Testing Service. (1978). *The final report for a project to develop twelve new forms of the Tests of General Educational Development and to standardize the tests nationally in the United States*. Princeton, NJ: ETS.

GED Testing Service. (1984a). *Examiner's manual for the Tests of General Educational Development*. Washington, DC: American Council on Education.

GED Testing Service. (1984b, September). *GED Items* (Vol. 1, No. 3). Washington, DC: American Council on Education.

Malizio, A. G., & Whitney, D. R. (1981). *Who takes the GED Tests?* (GEDTS Research Studies, No. 1). Washington, DC: American Council on Education.

Malizio, A. G., & Whitney, D. R. (1982). *Examinee and high school senior performance on the GED Tests: A detailed analysis*. (GEDTS Research Studies, No. 3). Washington, DC: American Council on Education.

Malizio, A. G., & Whitney, D. R. (1985, May). *Educational credentials in employment: A nationwide survey*. (OECC Research Brief). Washington, DC: American Council on Education.

Patience, W. M., & Whitney, D. R. (1982). *What do the GED Tests measure?* Washington, DC: American Council on Education.

Patience, W. M., & Whitney, D. R. (1983, April). *Reliability of item reviews for content validity, item quality, and bias*. Paper presented at the meeting of the American Educational Research Association, Montreal.

Reed, N. V. (1984, December). *Maryland GED graduates: Changes in education, employment, and perceptions of self-worth two years later*. Maryland State Department of Education, Division of Instruction, Adult and Community Education.

Sonnenblick, C. (1980). *High school equivalency candidate performance on the New York state high school reading competency tests and on the GED Tests*. (GEDTS Research Brief No. 2). Washington, DC: American Council on Education.

Spille, H. A. (1981a, July). *The 1980 norming of the GED Tests*. (GEDTS Memo #15). Washington, DC: American Council on Education.

Spille, H. A. (1981b, October). *Interpreting GED Test results*. (GEDTS Memo #21). Washington, DC: American Council on Education.

Spille, H. A. (1982, March). *Interpreting the GED scores of college applicants who have earned a high school equivalency credential*. (GEDTS Memo #23). Washington, DC: American Council on Education.

Spille, H. A. (1984, August). *Bibliography for the Tests of General Educational Development*. (GEDTS Memo #33). Washington, DC: American Council on Education.

Stoker, H. (1985, February). *Performance characteristics of Florida's early exit equivalency candidates*. A report to the Florida Department of Education, Bureau of Adult and Community Education. (Available from the GED office of the Bureau of Adult and Community Education, Florida Department of Education, Tallahassee, FL 32303.)

Wilson, R. C., & others. (1981, April). *The success of high school diploma and GED equivalency students in vocational programs at Lake City Community College, Florida*. (GEDTS Research Brief No. 4). Washington, DC: American Council on Education.

Wolf, J. C. (1980, March). *Predictive validity of the GED Tests for two-year college study: South Plains College, Texas*. (GEDTS Research Brief No. 1). Washington, DC: American Council on Education.

GED Research Briefs contain summaries of completed research studies relating to the development, use, and interpretation of results from the Tests of General Educational Development. *Briefs* are prepared by the authors and distributed by the GED Testing Service in the interest of improving the understanding and use of GED test results. *Briefs* may be reproduced for distribution to interested persons.

AMERICAN COUNCIL ON EDUCATION

One Dupont Circle • Washington, D. C. 20036